# Study of BGP Peering Session Attacks and Their Impacts on Routing Performance

Kotikalapudi Sriram, *Fellow, IEEE*, Doug Montgomery, *Member, IEEE*, Oliver Borchert, Okhee Kim, and D. Richard Kuhn, *Senior Member, IEEE*

***Abstract*—We present a detailed study of the potential impact of border gateway protocol peering session attacks and the resulting exploitation of route flap damping (RFD) that cause network-wide routing disruptions. We consider canonical grid as well as down-sampled realistic autonomous system (AS) topologies and address the impact of various typical service provider routing policies. Our modeling focuses on three dimensions of routing performance sensitivity: 1) protocol aware attacks (e.g., tuned to RFD); 2) route selection policy; and 3) attack-region topology. Analytical results provide insights into the nature of the problem and potential impact of the attacks. Detailed packet-level simulation results complement the analytical models and provide many additional insights into specific protocol interactions and timing issues. Finally, we quantify the potential effect of the BGP graceful restart mechanism as a partial mitigation of the BGP vulnerability to peering session attacks.**

***Index Terms*—Border gateway protocol (BGP), BGP graceful restart, BGP security, Internet routing protocol security, performance modeling, realistic topology, route flap damping (RFD), routing policy.**

## I. Introduction

THERE IS A growing apprehension in governments and the Internet industry that there are potentially significant vulnerabilities [1]–[17] in the deployed border gateway protocol (BGP) routing system [18], [19]. While to date there have been few, if any, serious focused attacks on the BGP infrastructure, researchers speculate and debate the potential of targeted attacks to trigger large scale, potentially cascading, failures and persistent instability in the global routing system [2]–[12]. In response to this situation, numerous proposals have been developed that attempt to provide varying levels of protection and assurance to various aspects of BGP's operation [20]–[23]. Each of these proposals implicitly embodies a somewhat different view of the attributes of the problem space and the practical constraints of the solution space. Unfortunately, the lack of a shared understanding of both the risks associated with focused attacks and the cost-benefit tradeoffs of various mitigation techniques will likely doom prospects for the rapid development and widespread adoption of a comprehensive set of solutions. It may also be noted that there are some efforts to even fundamentally rethink the design of BGP and the control plane design from a security point of view [24], [25].

Manuscript received September 15, 2005; revised March 30, 2006 This work was supported in part by the Department of Homeland Security under the Secure Protocols for the Routing Infrastructure (SPRI) Program and in part by the NIST Information Technology Laboratory Trustworthy Networking Program.

The authors are with the National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899 USA (e-mail: ksriram@nist.gov; dougm@nist.gov; okim@nist.gov).

Digital Object Identifier 10.1109/JSAC.2006.877218

To date, most modeling and analysis of BGP behavior under threatening scenarios has focused on postmortem analysis of global routing tables during worm and virus attacks of Internet hosts [26]–[28]. Unfortunately (or fortunately, depending upon one's perspective), there are no known live data or traces from large-scale attacks that were targeted at BGP itself. In order to fill this void, we have developed a simulation capability to model large scale attacks specifically focused on the BGP infrastructure. Our goal is to conduct "what if" analyses of yet unseen attacks and to develop means to characterize the impact of various attacks on a distributed BGP routing system. Of particular interest is the discovery of potential global emergent behaviors (e.g., cascading failures, persistent oscillations, permanently degraded routing) induced by successful local attacks, and the identification and evaluation of new BGP threat scenarios. We have extended the Scalable Simulation Framework Network (SSFNet) BGP simulation modeling tools [29], [30] to include an attack-modeling framework capable of generating arbitrary attacks with parameterized form, intensity, behavior, extent, and duration. In addition, we have developed metrics and an attack analysis framework capable of characterizing the impact of successful attacks in terms of their effects on global routing and the detailed operation of the BGP protocol [16], [31].

Our simulation tool has the capability to simulate several hundreds of autonomous systems (ASs). The AS-level topology can be a canonical grid or mesh, or a down-sampled realistic topology. In this paper, we study the impact of focused BGP peering session attacks and present simulation and analytical results. Through these results, it is revealed that malicious attackers could exploit route flap damping (RFD) mechanisms to amplify the duration of AS-to-AS or AS-to-prefix isolations. RFD is a method for receiver-side route monitoring and suppression in the event of frequent updates [32], [33]. However, that benefit has the flipside in that by sustained peering session attacks into various BGP sessions in an AS-path or into a portion of a network with many AS paths, attackers can cause isolation of ASs at the two ends of the attack region. We show that this potentially is a serious type of denial of service (DOS) attack, which is amplified by the particulars of BGP behavior (namely, RFD tuning parameters), and present a detailed quantitative analysis of its impact. Another dimension of our study is to incorporate realistic route selection policies. We quantify the impacts of attacks on BGP performance under three scenarios: one involving no policy and two increasingly restrictive policies, which are based on service provider relationships. Our study also develops an understanding of the sensitivity of routing performance to the topological focus of the attack region. The three

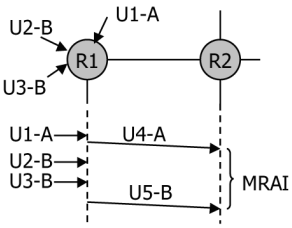


Fig. 1. The role of MRAI in update propagation.

areas of routing performance sensitivity, when subjected to session attacks, addressed here are: 1) protocol aware attacks (e.g., tuned to RFD); 2) route selection policy; and 3) attack-region topology. Our analytical results provide insights into the nature of the problem and impact of the attacks. Detailed packet-level simulation results complement the analytical results and provide many useful insights as well. We also quantify the effect of BGP graceful restart (BGP-GR) [34] mechanism on partial mitigation of the RFD-based BGP vulnerability.

In the rest of this paper, Section II presents an understanding of how random peering session attacks may trigger the RFD penalty and cause routes to enter the RFD suppression state. In Sections III and IV, the analytical model of RFD behavior during BGP session attacks and numerical results based on the analysis are presented, respectively. An analysis of the benefits of using BGP-GR mechanism is presented in Section V. Our SSFNet-BGP-based attack simulation framework and models are described in Section VI. Section VII deals with simulation results and their discussion.

## II. BGP Attacks With Exploitation of Route Flap Damping

We start here by providing brief introductions to the principles of BGP minimum route advertisement interval (MRAI) and the RFD. These collectively play a role in the models we develop to characterize the impact of peering sessions attacks and concomitant RFD exploitation. MRAI is a sender-side peering discipline designed to control the BGP update-processing load. Values of MRAI are randomly chosen in the range of 22.5–30 s on per peer basis. In the example in Fig. 1, router R1 receives in quick succession three BGP updates: U1-A about prefix A, and U2-B and U3-B about prefix B from different peers. These updates may arrive temporally close to each other at R1, but the MRAI at R1 causes them to be coalesced into fewer updates and/or separated from one another by at least an MRAI in their propagation to peer router R2. After processing, R1 sends update U4-A about prefix A to R2. Further, R1's route computation coalesces U2-B and U3-B into as single update U5-B, which is sent to peer R2 separated by an MRAI from U4-A. This example illustrates that the quantity and rate of updates can be potentially reduced due to MRAI.

RFD is a method for receiver-side route monitoring and suppression of oscillations or unstable paths. An upstream router assigns an incremental RFD penalty to a peer and destination (i.e., prefix) combination each time it receives a BGP update pertaining to that combination. If the RFD penalty exceeds a preset cutoff threshold, then the route is suppressed and withdrawals are sent to neighbors about the prefix in question. The RFD penalty is allowed to decay exponentially with a chosen halftime (i.e., decay constant). When it drops below a chosen reuse threshold, then the route is no longer suppressed and updates are once again processed for the peer-prefix combination in question. Table I shows the values of various RFD parameters for two common commercial implementations labeled as vendors A and B; the two sets of numbers are later used in this paper for a sensitivity study relative to RFD parameters.

TABLE I
RFD Parameter Values

| RFD Parameter | Vendor A | Vendor B |
|---|---|---|
| Withdrawal penalty | 1000 | 1000 |
| Re-advertisement penalty | 0 | 1000 |
| Attribute change penalty | 500 | 500 |
| Cutoff threshold | 2000 | 3000 |
| Half time | 900s | 900s |
| Reuse threshold | 750 | 750 |
| Max suppress time | 3600s | 3600s |
| Max penalty | 12000 | 12000 |

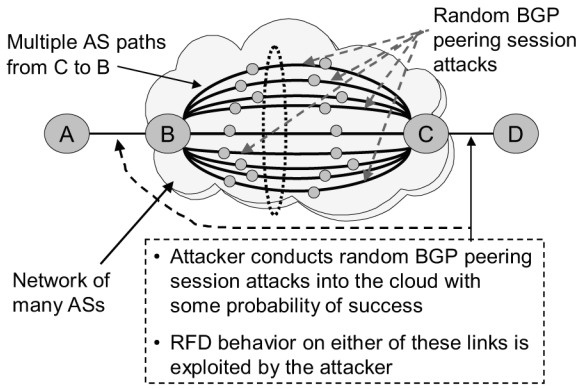


Fig. 2. Illustration of random BGP peering session attacks.

There are many different attack possibilities on the BGP routing infrastructure; an enumeration of BGP attacks is provided in [17]. We focus on attacks that cause the BGP peering sessions to be reset. A common way to reset a BGP peering session is to reset or attack the underlying transmission control protocol (TCP) connection. There are several known vulnerabilities associated with TCP and the Internet control message protocol (ICMP), which could be exploited to cause TCP connection-reset attacks [11]–[15]. One example is the "slipping in the window" TCP reset attack, which received a lot of attention recently [14], [15]. The success of this attack depends on the attacker's ability to correctly guess a TCP sequence number within a TCP flow control window. Spoofed ICMP error messages to cause TCP reset have also been brought to attention recently [11]. The ICMP-based attacks causing TCP resets do not require guessing the TCP sequence number. Hard or soft ICMP error messages can be potentially spoofed to cause TCP resets. The details regarding ICMP attacks against TCP are discussed in [11]. Fig. 2 illustrates how random BGP peering session attacks can lead to exploitation of the RFD by attackers, and cause prolonged AS-prefix isolations. Here, each node in the network represents an AS, as well as a prefix (i.e., destination). The figure illustrates that there may be multiple AS paths between ASs B and C. What is in the cloud is a network of many ASs. The attackers are assumed to have some

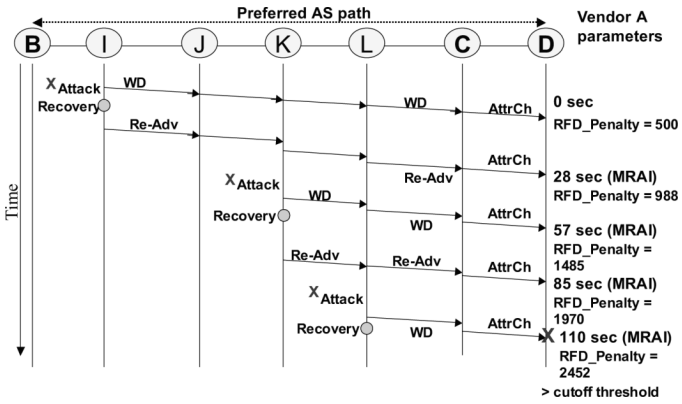


Fig. 3. Illustration of update message propagation and RFD penalty accumulation for the preferred AS path between nodes B and D.

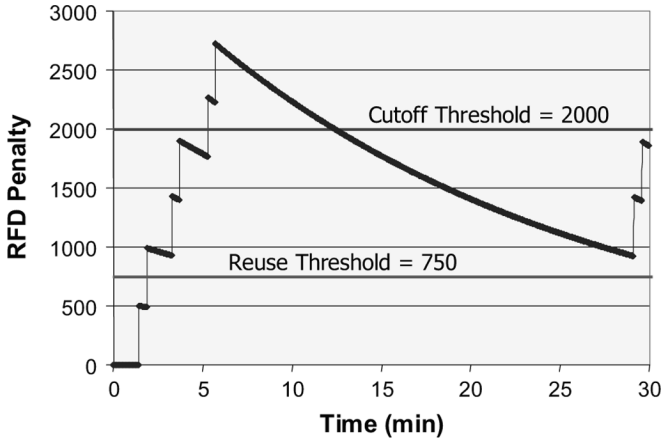


Fig. 4. Illustration of RFD penalty-based cutoff and recovery.

capability available that they use to launch BGP peering session attacks into the network, and depending on their method and collective resources, there is a measurable probability of success for each peering session attack. A successful attack would cause a peering session to be terminated, and cause the affected ASs to send withdrawals about all the prefixes in their routing tables that are rendered unreachable. The RFD behavior on links A-B and C-D would be exploited due to the attacks, and major outages (isolations) can result due to: 1) D imposing RFD suppression on (C,B) peer-prefix combination and all prefixes reachable via B and 2) likewise for A and (B,C).

The details of how this suppression works are further explained with the help of Fig. 3, where a linear topology (representing a single AS path), consistent with the individual alternate AS paths between B and D in Fig. 2, is considered. In Fig. 3, the progression of BGP updates horizontally from left to right is in the spatial dimension (hop to hop), and vertically from top to bottom is the progression of time. The figure shows three BGP peering session attacks happening on three different hops (B-I, J-K, K-L) in the AS path at different times. It shows the flow of updates, classified as either withdrawals (WD), re-advertisements (Re-Adv), or attribute-change (AttrCh). The BGP nodes along the way cause the updates to be separated by MRAI intervals. It is assumed that the peering session that was attacked is able to recover within a short time as compared with the MRAI. This quick recovery can occur, for example, when a BGP session is forced to terminate by a TCP reset attack. The BGP session is automatically reestablished immediately after the TCP connection is restored between the affected peers. An attack causes a withdrawal to be sent to neighbors and a recovery causes a Re-Adv to be sent. When C receives a withdrawal about B, if an alternate path is available (see Fig. 2), then C sends an AttrCh update to D informing D of an alternate route to B. When C receives a Re-Adv about B, it reverts to the previous path, and again sends an AttrCh update to D informing D of the reversal to the previous path. While these updates about B from peer C are received at D, the RFD penalty for B at D via peer C increases, and exceeds the cutoff threshold following just three attacks, as shown in the Figs. 3 and 4. This essentially isolates D from B and all prefixes reachable via B. In most implementations of BGP, if another attack is launched along the AS path before the RFD penalty reaches below reuse threshold, then the suppression (and isolation) continues even longer (see Fig. 4). When an update is regarded as a flap, the RFD penalty will be incremented due to the update even when the RFD is in a decay mode. Thus, attackers can effectively tune the attack rate to be at intervals roughly equal to MRAI or longer, trigger AS-prefix isolation within minutes, and sustain the isolation for long periods with a much slower rate of additional attacks.

## III. Analytical Model for Peering Session Attacks Triggering RFD Cutoff

The purpose of the analytical model presented here is to predict the probability of AS-prefix isolation under suitable assumptions regarding the attack characteristics. As described earlier, AS-prefix isolations are the result of the RFD penalty exceeding the cutoff on all alternative paths between the AS and the prefix. Here, we assume that the attacks happen independently on any of the BGP peering sessions in the cloud of Fig. 2. Thus, chances are that the RFD penalty will exceed the cutoff on different AS paths between BGP routers B and C in quick succession of one another. Thus, it is a reasonable approximation if we derive the probability of RFD penalty exceeding the cutoff for the longest AS path between B and C, and approximate that to the AS-prefix isolation probability in question. It can be reasoned that this would be in fact a good and slightly conservative approximation.

As shown in Fig. 5, we model the AS path between the endpoints of interest as $n-1$ BGP peering sessions (BGP router 1 to BGP router $n$). For purposes of modeling, we assume (without loss of generality) that the peering session between BGP routers $n$ and $n+1$ does not come under attack. The RFD penalty at BGP router $n+1$ is to be modeled in order to determine the probability of isolation at BGP router $n+1$ in relation to the peer router $n$, destination router 1 and prefixes reachable via 1. Attacks spaced closer than MRAI interval do not speed up the time to isolation, and hence it is meaningful to assume that the attacks would be spaced approximately at MRAI time intervals. We assume that a router's control plane may be compromised

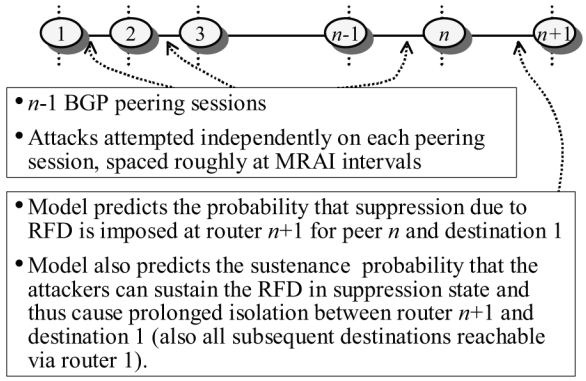


Fig. 5. Analytical model for AS-prefix isolation probability.

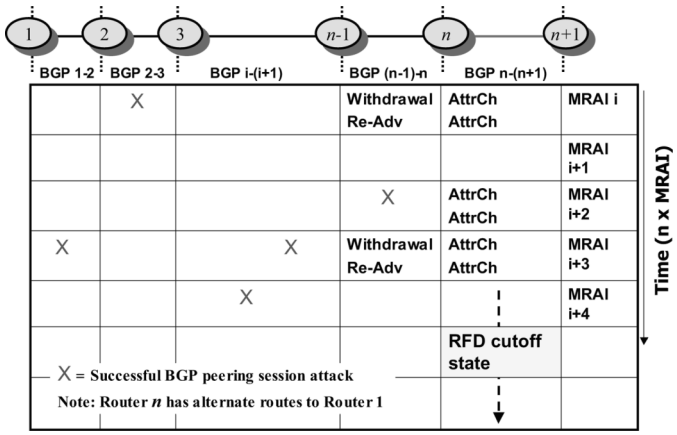


Fig. 6. Time–space model for relating RFD penalty accumulation to attacks and derivation of probability of successful AS-prefix isolation.

with probability $p$ and an associated BGP peering session may be then attacked with probability $q$. Thus the probability of a successful BGP peering session attack is $Q = pq$. The model we describe below derives the probability that route suppression happens due to RFD at BGP router $n+1$ for peer $n$ and destination 1. The model also predicts the probability that the attackers can sustain the RFD in suppression state, and thus cause prolonged isolation between router $n+1$ and destination 1.

The basic principle of the time–space model of the attacks and RFD penalty accumulation can be explained with the help of Fig. 6. In the time–space matrix illustrated in Fig. 6, each cell represents a hop (or BGP peering session) location and the time in multiples of MRAI. Although the MRAI is a variable in the range 22.5–30 s, here we assume it be a constant with a fixed value of 26 s. The X's in the figure represent peering session attacks. Just to get a conservative estimate of how soon the RFD penalty could accumulate, we make two reasonable assumptions: 1) attacks occur at intervals approximately matched to MRAI and 2) when an attack happens, no other updates have recently happened within approximately an MRAI time so that the updates resulting from the attack in consideration propagate quickly across from left to right in Fig. 6 well within an MRAI time interval. The second assumption is aided in its accuracy partially due to the first assumption. To describe the stochastic model, let us define the following parameters:

| | |
|---|---|
| $C$ | cutoff threshold; |
| $R$ | reuse threshold; |
| $H$ | halftime (decay parameter); |
| $T$ | MRAI time; |
| $P$ | incremental RFD penalty incurred per successful attack event; |
| $n$ | number of BGP nodes in the AS path subject to attacks (see Fig. 6); |
| $Q$ | Prob. {a BGP peering session attack is successful}; |
| $\theta$ | Prob. {AS path of $n$ ASs is successfully attacked at one or more BGP peering sessions}; |
| $E$ | elapsed time from the time of beginning of BGP session attacks (in multiples of MRAI); |
| $R_P(n+1;n,1;iT)$ | RFD penalty at router $n+1$ for peer $n$ and destination 1 at time $iT$; |
| $\alpha(n,k)$ | Prob.$\{R_P(n+1;n,1;iT) > C$ for some $i \in (1,k) \vert E = kT\}$. |

In the above definitions, the incremental RFD penalty, $P$, incurred per successful attack is assumed to be one number even though Table I shows different penalty values for different types of updates. This is because in the system we are modeling, each BGP session attack eventually produces a pair of AttrCh updates between nodes C and D (see Fig. 3) or between nodes $n$ and $n+1$ in our analytical model in Fig. 6. These AttrCh updates are still decipherable by the receiving peer as corresponding to withdrawal (implicit) or Re-Adv. In effect, corresponding to each successful attack, the net incremental penalty $P$ will be $1000(= 1000 + 0)$ for the case of Vendor A and $2000(= 2{*}1000)$ for the case of Vendor B. The key performance metric of interest in this analysis is $\alpha(n,k)$, the probability that the attackers can cause RFD triggered AS-prefix isolations in $kT$ time interval or less.

From Fig. 6, given that there are $n$ BGP peering sessions in the AS-path of interest, it can be deduced that the probability $\theta$ of successful BGP attack on the $(n-1)$-hop AS path in an MRAI interval is given by

$$\theta = 1 - (1 - Q)^{n-1}. \tag{1}$$

The probability, $\beta_i(n,k)$, that there are $i$ successful attacks on the $(n-1)$-hop AS path in time $kT$ (equivalently, $k$ successive MRAI intervals), is given by

$$\beta_i(n,k) = \frac{k!}{i!(k-i)!}\theta^i(1-\theta)^{k-i}, \qquad i \le k. \tag{2}$$

Let us define $k_m$ as the absolute least number of successful BGP attacks needed on the AS-path in consideration to cause the RFD penalty to exceed the cutoff threshold, $C$. Then, we have

$$k_m = \left\lceil \frac{C}{P} \right\rceil. \tag{3}$$

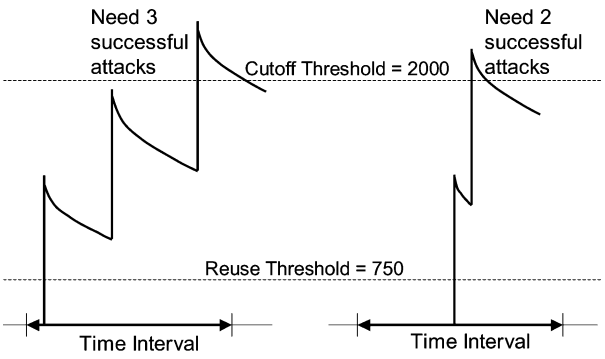


Fig. 7. Model for estimation of attacks needed to push penalty above cutoff.

Clearly, if $k_m$ consecutive MRAI intervals have successful peering session attacks, and if there were no exponential decay, then the cumulative penalty would meet or exceed the cutoff threshold $C$. However, in reality, the exponential decay of RFD penalty must be taken into account (see Fig. 7). If the attacks are bunched together (closely spaced), and located towards the beginning or the end of the $k$-MRAI interval, then the decay would be too much or too little, respectively, and would not lead to a realistic estimate of minimum number of attacks needed to exceed the cutoff. Thus, while taking the exponential decay into consideration, it is reasonable to assert that the attacks in the $k$-MRAI time period can be spread nearly evenly to derive a realistic estimate of the minimum number of attacks, $j_{\min}(k)$, needed to meet or exceed the cutoff threshold, $C$ (see Fig. 7). Thus, for a given $k$, the $j_{\min}(k)$ can be estimated by finding the smallest integer $j$ for which the following inequality is satisfied:

$$P\sum_{i=0}^{j-1} 2^{\{-\frac{ikT}{(j-1)H}\}} > C, \quad k \geq k_m. \tag{4}$$

Once $j_{\min}(k)$ is known, then the key performance of metric interest, $\alpha(n,k)$, is derived as follows:

$$\alpha(n,k) = \sum_{i=j_{\min}(k)}^{k} \beta_i(n,k), \quad k \geq k_m. \tag{5}$$

Based on available BGP protocol descriptions in the literature [32] and [33], it appears that the RFD specifications require the penalty be incremented even when the RFD is in a cutoff (suppression) state, provided that the received update is a flap. As a result, it is possible that, if another attack is launched along the AS path before the decaying RFD penalty reaches below the reuse threshold, then the RFD penalty may be incremented further until the maximum penalty value (12 000) is reached (see Table I). Thus, route suppression and AS-prefix isolation continue even longer (as previously noted in Fig. 4). We define the probability of sustenance $P_{\text{sus}}$, as the probability that the AS-prefix isolation, once reached, is sustained further by launching at least one additional successful peering session attack on the AS-path in consideration before the RFD penalty goes below the reuse threshold. This probability of sustenance, $P_{\text{sus}}$, is given by

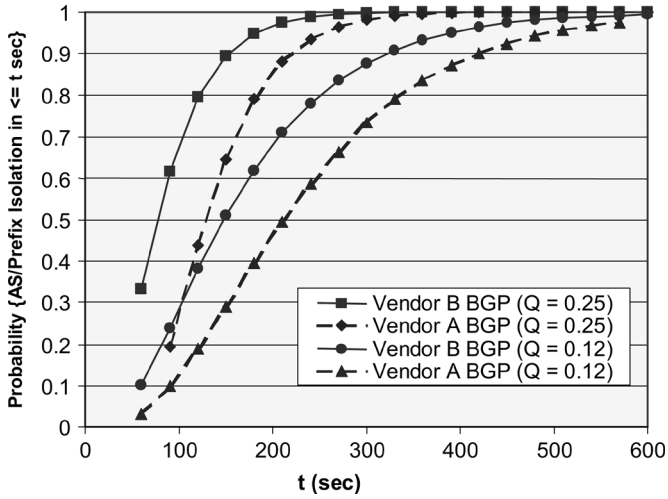


Fig. 8. Probability of AS-prefix isolation—sensitivity to vendor parameters and probability of successful session attack $(n = 4)$.

$$P_{\text{sus}} = 1 - (1-\theta)^{\lceil \frac{H(\log_2 \frac{C}{R})}{T} \rceil}. \tag{6}$$

This equation essentially estimates the probability that at least one successful attack can be launched on the AS-path in consideration during the decay time from the cutoff threshold $C$, to the reuse threshold $R$.

## IV. Numerical Results From Analytical Model

In this section, we present some numerical results based on the analytical model of the preceding section. The probability that AS-prefix isolation happens in time $t$ seconds or less is shown for the case of $n = 4$ hops in Fig. 8. In this plot, we also show the sensitivities to the vendor parameters, as well as to the probability of success of a peering attack $Q$. The performance (or vulnerability) is worse for Vendor B because the incremental penalty per successful attack (withdrawal plus Re-Adv) is much higher for Vendor B (2000) versus Vendor A (1000). This effect is dominating even though Vendor B has a higher cutoff threshold than Vendor A (see Table I). As would be expected, Fig. 8 also shows that higher probability of AS-prefix isolation occurs for higher values of $Q$.

With longer duration of attacks or larger area of vulnerability in the network, the attackers have a greater chance to be successful. This is illustrated in Fig. 9 where the probability of AS-prefix isolation goes higher as the number of hops $n$ in the AS-path increases. It is generally known that the typical AS-path length in the Internet is about 4. The three-dimensional (3-D) plot of Fig. 10 further illustrates how the probability of AS-prefix isolation increases with the number of hops, as well as the time duration of attacks. Fig. 11 shows the probability of sustenance $P_{\text{sus}}$, as a function of $Q$ [see (6)]. It illustrates that it is probabilistically much easier to sustain the AS-prefix

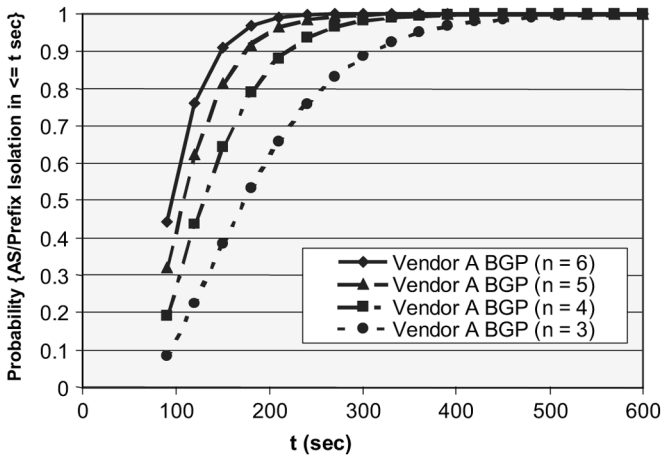


Fig. 9. Probability of AS-prefix isolation—sensitivity to vendor parameters and AS-path length ($Q = 0.25$).

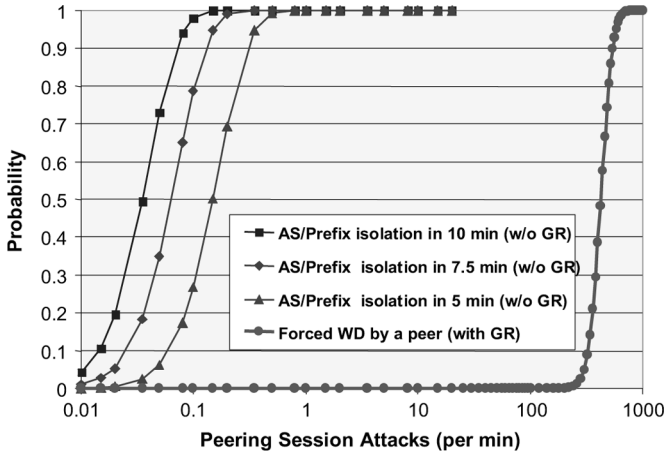


Fig. 12. Probability of adverse impact on routing as a function of peering session attack rate for BGP with/without GR ($Q = 0.25, n = 4$).

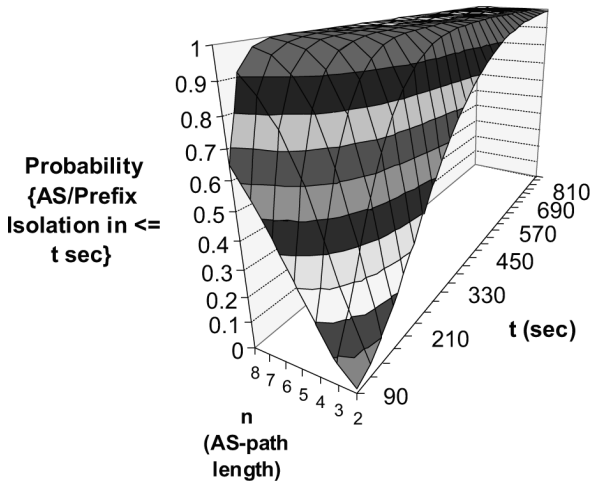


Fig. 10. Probability of AS-prefix isolation (3-D view) as a function of time and AS-path length ($Q = 0.25$).

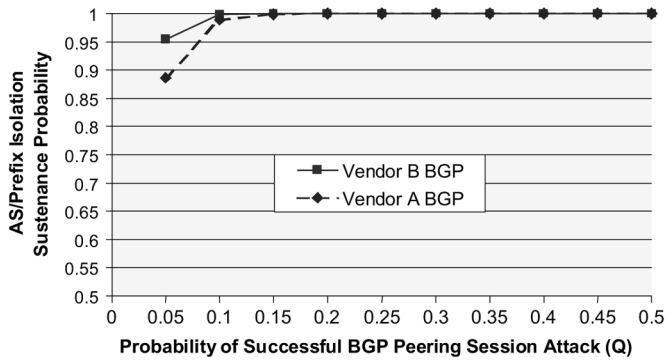


Fig. 11. Probability of sustenance of AS-prefix isolation as a function of success probability of peering session attack ($n = 4$).

isolation as compared with achieving the isolation initially [see explanation leading to derivation of (6)].

## V. Benefit of BGP Graceful Restart (BGP-GR)

The BGP-GR mechanism [34] gives a downed router (downed in the control plane only) time to restart without peers withdrawing its routes. This option is negotiated between peers at the time of BGP peering session establishment. Two flags bits used in capability advertisement during BGP-GR negotiations are: 1) restart bit—used to indicate if router has restarted and 2) forwarding bit—used to tell a peer router that the capability exists to preserve forwarding state through a restart period. Once a router has announced its BGP-GR capability, during its restart (of BGP or BGP peering session) its neighbors do not immediately delete routes via that peer so that undue route flapping is prevented. A restart timer is used at each peer to determine how long it would wait before deleting stale neighbor routes. If the BGP open message is not received from the restarting router before the expiry of the restart timer, then the restart is presumed failed, routes previously announced by that peer are deleted, and withdrawals are sent.

Without BGP-GR, it is expected that the peering session attacks exploiting RFD behavior would be much more feasible. BGP-GR helps mitigate the effects of this type of attack. We have extended the analysis of Section III to model the impact of peering session attacks when BGP-GR is used. The analysis of Section III has been slightly modified to start with a Poisson attack arrival model for the BGP peering session attacks. As previously stated, each attack would have a given probability of success. From these assumptions, the probability of a successful session attack $Q$ on a BGP peering session in an MRAI interval can be determined. Thus, $Q$ would now be a function of the attack arrival rate. Now, the equations in Section III can be readily used to derive performance metrics of interest as a function of the attack arrival rate.

To model the effect of BGP-GR, it is to be noted that BGP-GR allows a router's control plane (or BGP peering sessions) to be attacked many times over the duration of the restart timer without causing any disruptions in the forwarding plane. The attackers have to persistently attack within each BGP session recovery time over the duration of the restart timer in order to cause the peers to send withdrawals. We capture this in our analytical model of the attack effects with BGP-GR. The plots shown in Fig. 12 comparing the performance with and without BGP-GR are very instructive. We assume here that the BGP-GR restart timer is 120 s and the BGP session recovery time is about 4 s (much smaller in reality when TCP reset attack causes BGP

session closure). The three plots to the left in Fig. 12 show the probability that AS-prefix isolation can be achieved in 5, 7.5, and 10 min, respectively, at comparatively low rates of peering session attacks when BGP-GR is not used. In contrast, at least two orders of magnitude higher rates of attacks are required in order to even cause forced withdrawals by peers when BGP-GR is used.

There seem to be some practical concerns about use of BGP-GR. Most newer BGP routers in the service provider networks have BGP-GR capability but it is very rarely (if at all) turned on. A BGP best practices document from the NISCC notes that "Several providers (U.S.) suggest that the cost of implementing this feature [BGP-GR] outweighs the benefits" [15]. Based on an informal survey of some ISPs, we found that their customers seem to prefer, in the case of multihoming, that routing be done via a completely healthy BGP router (including control plane) rather than use BGP-GR over a route where the control plane is compromised. Their preference seems to be in that they wish to avoid a BGP router that is in recovery because it could have stale routing information in its forwarding information base (FIB).

## VI. Description of the Simulation Model

Our simulation environment is based on SSFNet [29]. SSFNet provides a set of modules to simulate traffic at the Internet protocol (IP) layer and above. To support our work, we made several extensions and modifications to the TCP/IP and BGP modules. The default version of SSFNet does not come with a complete TCP state machine. Here, code was added to produce proper failure messages and warnings, as well as the detection of 1/2 open TCP connections. This extension allowed the simulation of spoofed TCP session resets, which in turn simulates BGP session attacks.

The existing implementation of the BGP protocol in SSFNet also had to be extended. The set of BGP modules that come with the SSFNet distribution did not include the "uncontrolled shutdown" of a BGP sessions. This scenario occurs if a successful hostile attack on the underlying TCP session results in a breakdown of the transport connection. Here, the interaction between BGP and TCP, as well as the proper shutdown and restart mechanism within BGP had to be implemented. This included reinitializing the RFD penalties (for prefixes reached via each other) to zero at each of the two affected peers after BGP session shutdown and reestablished. We used the RFD implementation available in SSFNet BGP (with our aforementioned modification), which is based on RFC 2439 [32].

In addition to the extension of existing modules, we designed a BGP attack modeling framework. This framework allows the installation and configuration of individual "attacker" modules into each BGP router. All modules have some basic attributes in common, such as attack execution probability, module activation start time, and duration of attack activity. These parameters can be scripted for each AS or AS-group separately and/or for all ASs globally. For this study, we use a single type of attack module, namely, a TCP-session attacker. We are currently in the process of designing and running experiments with other types of attack modules, such as BGP message spoofing/tampering attacks and other protocol aware attacks (besides the one presented here). These are being documented elsewhere [31].

The following choices of BGP parameters and features are common to all simulation experiments in this study: 1) Vendor A's RFD parameters are used (see Table I); 2) MRAI is randomly chosen each time at each BGP router from a uniform distribution over the range 22.5–30 s; 3) if two AS-paths have equal cost in BGP route computation, then tie breaking based on lower IP address of peer AS is used; and 4) split horizon with poison reverse is used for loop prevention.

## VII. Simulation Results and Discussion

For packet-level (i.e., BGP message-level) simulations, we allow the network a startup and initial stabilization time of 500 s, by which time all routes have converged to their respective stable routes. The attacks are launched starting at 500 s, and go on for 500 s for each of the experiments (grid and down-sampled realistic topology cases). It may be noted that our observation interval is on the order of 1000's of seconds with temporal snapshots taken every 10 s. This is done so that the MRAI and RFD related ripple effects of the attacks are captured well, and the observation interval goes well beyond the time when these ripple effects settle down. When RFD penalty is triggered, the many route restorations to respective stable paths happen after 1000's of seconds following the attacks. Note that the halftime of the exponential decay of RFD penalty is 900 s. Hence, the observation time should be at least several 1000's of seconds.

In all our experiments, the nodes represent ASs, as well as destinations (i.e., prefixes). In the grid, as well as the realistic topology, each AS contains one destination or prefix.

We now proceed to present simulation results based on a canonical grid (Section VII-A) and a down-sampled realistic topology (Section VII-B). The route selection policy and attack-region topology considerations are described and incorporated only in the case of simulations with the down-sampled realistic topology in Section VII-B.

### A. Grid Topology

We first consider a canonical grid topology with no policy constraint, which offers many alternate paths between any two ASs. We used the following network and attack parameters: 1) $16 \times 16$ grid topology (256 BGP nodes or routers); 2) attacks can occur on any of the 144 BGP peering sessions associated with the center $8 \times 8$ subgrid; 3) attack duration is 500 s; 4) attack duration is divided into 50 intervals of 10 s each, and there is the potential for one BGP session attack per peering session in each interval; 5) the probability of success of each session attack is assumed to be 25%; and 6) the timing of an attack is uniformly random within each interval.

Here, we attempt to provide insights into the impacts due to the attacks in terms of several performance metrics related to route stability and degradation of route quality. In the results that follow, the metrics will be compared for the cases of (a) without RFD and (b) with RFD. The purpose of this comparison is to show how the attacks—when tuned to protocol specifics (RFD,

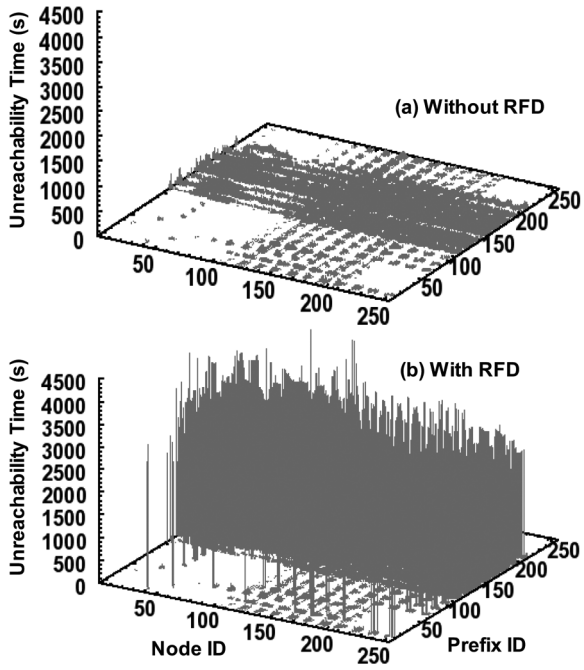


Fig. 13. Unreachability times for node-prefix pairs: (a) versus (b) comparison highlights amplification attributable to RFD.

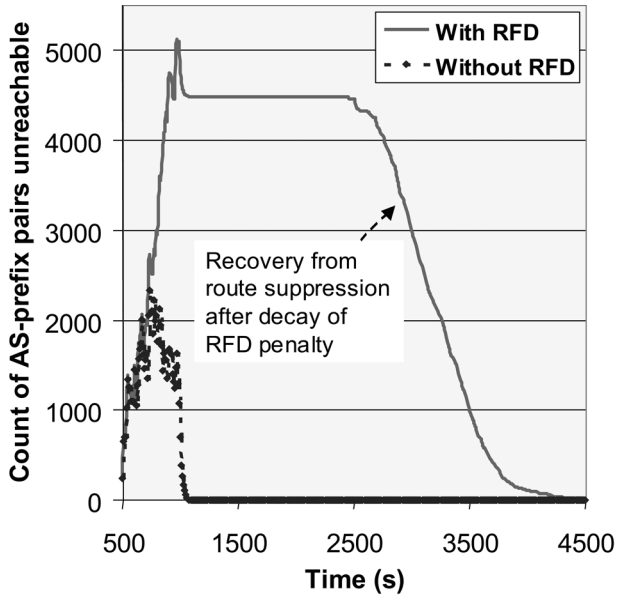


Fig. 14. Count of AS-prefix pairs unreachable: comparison shows that unreachability lasts about eight times longer for the case with RFD.

MRAI), can cause much more amplified damage to the routing infrastructure than otherwise.

For the purpose of this study, the unreachability time between a node and a prefix is defined as the total time (summed over possibly multiple occurrences) that the prefix is unreachable from the node during an experiment's observation interval. In Fig. 13, the unreachability times are plotted in 3-D as a function of the node and prefix IDs, and a comparison is shown between two cases: (a) without use of RFD and (b) with use of RFD. Because of the intense nature of the attacks, many prefixes are rendered unreachable from almost any node for some duration of time during the observation period. What is most striking in Fig. 13 is the fact that the node-prefix unreachability time in the case with RFD is typically larger by more than an order of magnitude as compared with that in the case without RFD (approximately 4500 and 350 s for the cases with RFD and without RFD, respectively). While RFD serves its purpose in terms of damping undesirable route flaps during normal operation of the Internet, it could however, aid the malicious attackers in terms of amplifying the impact of focused attacks.

This observation is further reinforced when we look at another important metric in Fig. 14. The count of AS-prefix pairs unreachable is a very useful routing performance metric since it tells us the total number of such pairs where the prefix in each pair is unreachable from the corresponding AS at a given time. The span of the $x$ axis (the time span) before the count of AS-prefix pairs unreachable goes down to zero is important to note in interpreting Fig. 14. This time span is about 560 s in the case without RFD, while the same is about 4000 s in the case with use of RFD. The attacks last for 500 s. Without RFD it takes about 60 s (1060 s $-$500 s $-$500 s) after that for all the routes to converge back to reachable routes (mostly stable routes), while it takes about 3500 s (4500 s $-$500 s $-$500 s) for the same to happen when RFD is in use. In the case of BGP without RFD, at most 2339 AS-prefix pairs are rendered unreachable by the attacks, all recovering quickly after the attacks subside. By comparison, in the case of BGP with RFD, about 4500 AS-prefix pairs are rendered unreachable, lasting for a prolonged period even after the attacks subside.

The reason for the plateau in the recovery of routes in Fig. 14 can be explained as follows. We observed that when the RFD penalty exceeds the cutoff threshold of 2000, in many cases it exceeds the same by approximately a withdrawal penalty of 1000. Thus, for many of RFD affected AS-prefix pairs when the RFD exponential decay begins after a cutoff, it starts to decay from a value just under 3000. From there, it takes about 1800 s to decay to the value of the RFD reuse threshold (=750). Hence, a large number of AS-prefix pairs see an unreachability duration of 1800 s and other pairs experience even larger durations of unreachability. This gives rise to the AS-prefix unreachability plateau in Fig. 14, followed by recovery to their respective stable paths. The isolation of a majority of AS-prefix pairs due to RFD cutoff begins at various times in the 500–1000 s attack interval, and the recoveries occur gradually in the approximate interval of 2500–6000 s depending on the accumulated values RFD penalty and their binary-exponential decay.

Fig. 15 shows the BGP update count versus time for cases (a) without RFD and (b) with RFD. The update count is the number of updates seen network-wide, collected over 10 s intervals. The update counts are classified as advertisements and withdrawals. It can be observed that the update activity in Fig. 15 temporally correlates with the rise and fall of number of AS-prefix pairs unreachable in Fig. 14. When session attacks are launched on many of the BGP sessions in the network over an attack interval of 500 s (500–1000 s), a sharp rise follows in the number of updates and also in the number of AS-prefix pairs unreachable. The second flurry of updates after the initial big flurry in Fig. 15(b) can be again explained with the RFD

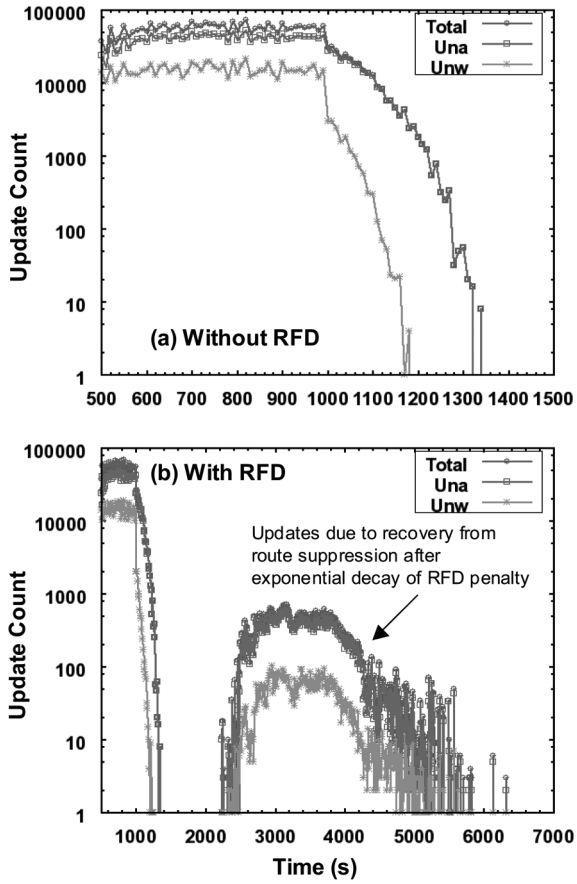


Fig. 15. BGP update count as a function of time: (a) versus (b) comparison shows the suppression followed by flurry of updates after RFD penalty decays.

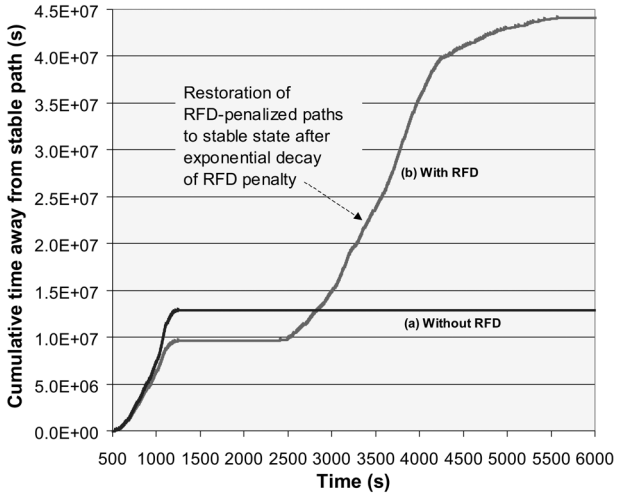


Fig. 16. Cumulative route deviation time (away from stable path): comparison shows deviations last much longer when RFD is used.

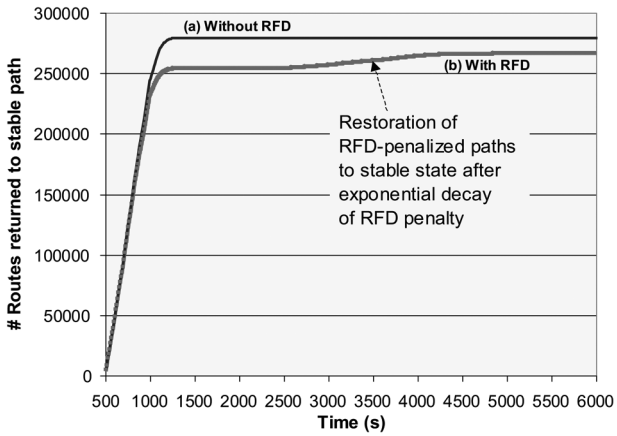


Fig. 17. Cumulative count of route returns to stable path.

exponential decay arguments (as was done explaining Fig. 14). These updates in the 2200–6000 s time interval correspond to the gradual recovery of RFD suppressed paths; the update announcements dominate over withdrawals. It may also be noted that a portion of the updates in the second flurry was observed to be explicit withdrawals, which result when routes are restored to stable paths and the split-horizon (with poison reverse) mechanism generates necessary withdrawals to avoid loops. The updates subside in about 800 s from the onset of attacks in the case of without RFD versus about 5500 s in the case with RFD. Looking at the details in the first 800 s in Fig. 15(a), it can be noted that the periodic mini-spikes in the update count happen at intervals of about 20–30 s. This corresponds to the MRAI, which randomly varies from 22.5–30 s. MRAI causes bunching to occur in the propagation of updates in the network. These MRAI influenced mini-spikes are seen for both advertisements, as well as withdrawals.

Two other routing performance metrics of interest are related to route deviations from the stable routes and can be regarded as route quality metrics. These are: 1) the cumulative time away from the stable paths over all routes network-wide and 2) the cumulative number of routes that have returned to their respective stable paths network-wide. In Figs. 16 and 17, these metrics are plotted versus time, and compared for the cases of (a) without RFD and (b) with RFD. At the end of the observation interval in Fig. 16, the cumulative time away from the stable paths is higher by a factor of 3.5 for the case with RFD ($\cong$4.5 W $10^7$ s) as compared with that for the case without RFD ($\cong$1.3 W $10^7$ s). Now, looking at Fig. 17, fewer route deviations are observed in the case with RFD because some of the routes have already been suppressed shortly after the onset of attacks, and hence are not subject to deviations. However, the suppressed routes wait through the long exponential RFD decay period before they are restored to their stable paths. In the case without RFD, all affected routes return to stable paths within tens of seconds after the attacks subside. However, the same takes thousands of seconds in the case with RFD.

### *B. Down-Sampled Realistic Topology*

Now, we proceed to present results based on a down-sampled realistic AS-level topology. Here, we also focus on the effects that service provider policies for route selection may have on routing performance under peering session attacks. Detailed and useful data about the AS-level topology of the Internet are available from the UCLA Internet Research Laboratory [35]. We have developed algorithms that work with this detailed data to down-sample and prune the very large topology (about 23 000 ASs and 96 000 BGP peering links) to a smaller

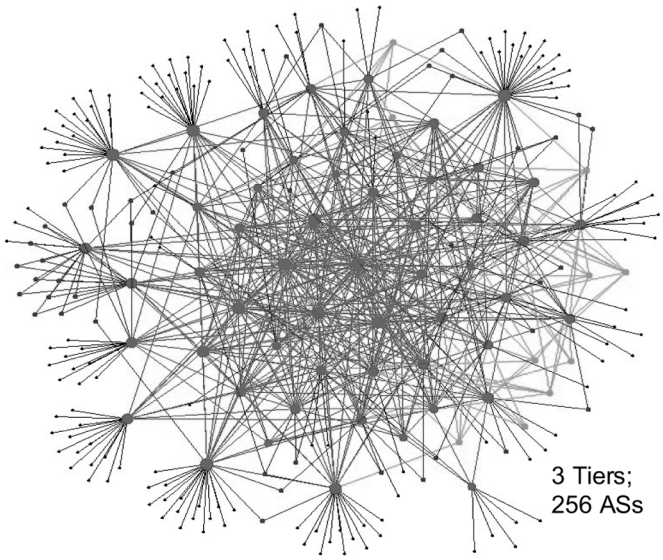


Fig. 18. Down-sampled realistic AS level topology in 2-D view.

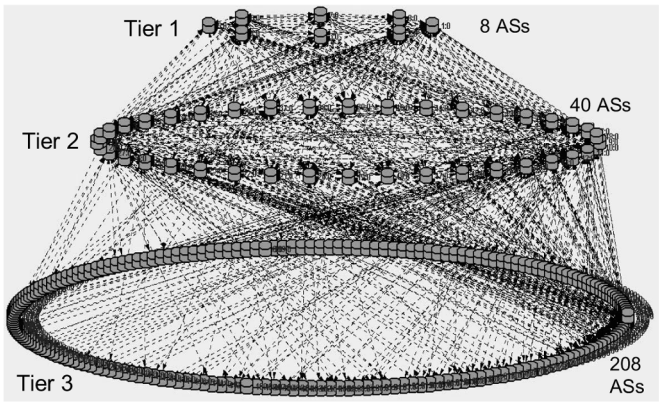


Fig. 19. Down-sampled realistic AS level topology in 3-D view.

TABLE II
PROPERTIES OF THE DOWN-SAMPLED REALISTIC TOPOLOGY

| | # ASs | # Peering Links | Avg. # Peering Links per AS |
|---|---|---|---|
| Tier 1 | 8 | | |
| Tier 2 | 40 | | |
| Tier 3 | 208 | | |
| Tier 1-1 | | 27 | 6.8 |
| Tier 1-2 | | 214 | 26.8 |
| Tier 2-2 | | 100 | 5 |
| Tier 2-3 | | 412 | 10.3 |
| Tier 3-3 | | 0 | 0 |
| Network-wide | 256 | 753 | 5.9 |

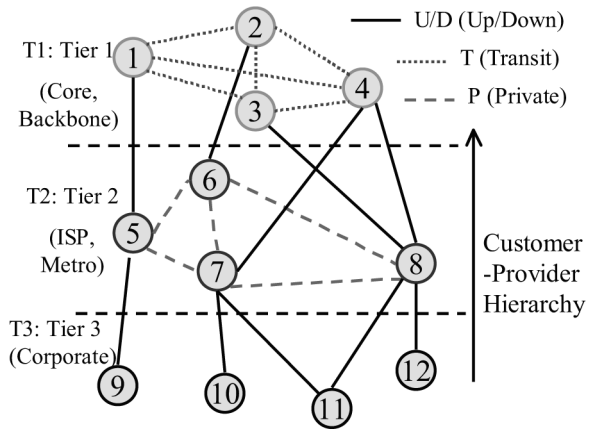


Fig. 20. Example AS-level network for policy illustration.

topology that can be handled in our SSF-based BGP simulations. The details of these algorithms are discussed in [31]. The algorithms inherently guarantee that the down-sampling of ASs and pruning of peering links are done in such a way that: 1) there are no disjoint networks in the reduced topology and 2) each AS in a lower tier is connected to at least one AS in the tier immediately above it. With the use of these algorithms, we obtained a down-sampled realistic AS-level topology that consists of 256 ASs and 753 peering links. Two illustrations of this topology are shown in Figs. 18 and 19 in two-dimensional (2-D) and 3-D visualizations, respectively. Our algorithm also assigns tier levels to ASs from the highest to the lowest tier levels based on the richness of ASs' peering connectivity. Table II shows the details of this topology in terms of numbers of ASs and peering links within each tier, as well as between tiers. The numbers of ASs in Tiers 1, 2, and 3 are 8, 40, and 208, respectively. The network is almost a full mesh for the Tier 1 ASs. Each Tier 1 AS is peered with 26.8 Tier 2 ASs on average. Each Tier 2 AS is peered with 5.4 Tier 1 ASs and 5 Tier 2 ASs on average. Thus, there is very rich peering connectivity at the two top tiers of the topology. Each Tier 2 AS is peered with about ten Tier 3 ASs (customers) on average. Each Tier 3 AS has two Tier 2 peers on average. Many of the Tier 3 ASs are stub nodes. Overall, the network-wide average for the peering connectivity is 5.9.

We consider two route selection policies based on service provider peering relationships. As illustrated in Fig. 20, the peering relationship between ASs of different Internet service providers (ISPs) (within a Tier) may be designated as transit (T) or private (P). Also, illustrated in Fig. 20 are U/D (up/down) links, which carry traffic up and down between tiers. A transit link allows carriage of any transit traffic. Generally, Tier 1 (i.e., core) ASs serve as transit points and their peering links carry transit traffic for the tiers below in an unrestrictive manner. Traffic that originated in any tier is allowed to pass over multiple hops in Tier 1 before it terminates at a Tier 1 AS or goes down to a Tier 2 AS. However, Tier 2 ISPs may be regional, and hence more restrictive in terms of their routing policy. They may designate their peering links as private (P) and allow for carriage of transit traffic only between their direct customers in the tier directly below. Keeping these possible peering arrangements in mind, we consider two route selection policies, as described in Table III. In Policy 1, all links within each tier are assumed to be type T and any number of transit hops are permitted within a tier for the traffic coming from a tier above or below or from within the same tier. However, the only restriction in Policy 1 is that once the traffic utilizes a D link, it cannot traverse a U link after that. This is specified by the $[U|T]*[D|T]*$ formula, where $|$ indicates "OR" function and $*$ indicates unrestrictive number of usages. Policy 2 on the other hand is much more restrictive. As described in Table III, in Policy 2 it is assumed that all links in Tier 2 are P links.

TABLE III
ROUTE SELECTION POLICY SPECIFICATIONS

| | Route Selection Rule | Peering Link Designations |
|---|---|---|
| Policy 1 | [U\|T]*[D\|T]* | All intra-Tier links are Transit (T) |
| Policy 2 | {[U\|T]*[D\|T]*} OR {[U]*[P]?[D]*} | All intra-Tier links in Tier 1 are Transit (T); All intra-Tier links in Tier 2 are Private (P) |

TABLE IV
LIST OF EXPERIMENTS WITH DOWN-SAMPLED REALISTIC TOPOLOGY

| Experiment | Attack region | Policy |
|---|---|---|
| E1P0 | All links subjected to attacks | no Policy |
| E1P1 | - do - | Policy 1 |
| E1P2 | - do - | Policy 2 |
| E2P0 | T1-T1 and T1-T2 links subjected to attacks | no Policy |
| E2P1 | - do - | Policy 1 |
| E2P2 | - do - | Policy 2 |
| E3P0 | Only T2-T3 links subjected to attacks | no Policy |
| E3P1 | - do - | Policy 1 |
| E3P2 | - do - | Policy 2 |

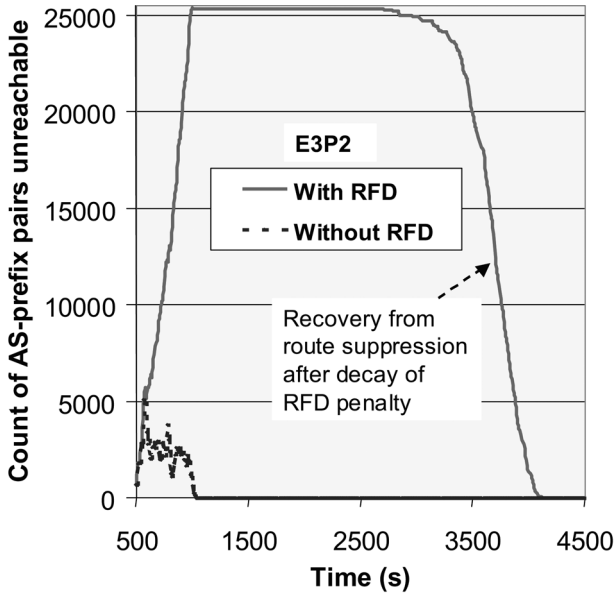


Fig. 21. Count of AS-prefix pairs unreachable as a function of time: comparison shows that unreachability gets much worse due to RFD.

Further, each of these P links only allows transit for data of the customer ASs in Tier 3 that are directly connected to one of the two peering ASs that share the P link. Thus, in Policy 2, once traffic traverses a P link, it can only use D link(s) after that. This is specified by the general formula $\{[U|T]*[D|T]*\}$ OR $\{[U]*[P]?[D]*\}$, where ? indicates zero or one use. The first principle for route selection is still lowest hop-count (i.e., highest linkpref value) but the candidate routes must meet the policy restrictions additionally. The number of feasible routes for an AS-prefix pair will be in decreasing order from the case of no policy to Policy 1 to Policy 2.

For the simulation experiments, the basic attack parameters are as follow: 1) attack duration is 500 s; 2) attack duration is divided into 50 intervals of 10 s each, and there is the potential for one attack per peering session in each interval; 3) the probability of success of each peering session attack is assumed to be 25%; and 4) the timing of the attack is uniformly random within each interval. The topology of the attack region is varied as described below. The set of experiments we performed with the down-sampled realistic topology are listed in Table IV. There are essentially three sets of experiments based on the attack-region topology: 1) all links: all peering links are subject to session attacks; 2) T1-T1 and T1-T2 links: only the Tier 1 to Tier 1 and the Tier 1 to Tier 2 peering links are subject to session attacks, and 3) T2-T3 links: only the Tier 2 to Tier 3 peering links are subject to session attacks. Each of these sets are further classified by policy: P0 represents no policy is used, P1 represents Policy 1 is used and P2 represents Policy 2 is used. There are nine combinations of attack region and policy. Accordingly, there are nine experiments enumerated in Table IV.

The BGP performance metrics that were considered for the grid topology are also used here. Path-length is an additional metric considered here. First, we present a result related to the amplification of unreachability due to RFD. For this, we consider experiment E3P2 (attacks on T2-T3 links and Policy 2). Fig. 21 shows the count of AS-prefix pairs unreachable over time, comparing the cases of BGP with RFD and without. We note that at most about 5000 AS-prefix pairs are unreachable for the case without RFD while about 25300 AS-prefix pairs are unreachable for the case with RFD. The comparison is even worse if we also take into consideration the maximum time-spans of this unreachability, which are about 500 s for the case without RFD versus about 3500 s for the case with RFD. These observations regarding amplification of unreachability due to RFD are much more pronounced for the case of realistic topology as compared with those for the grid topology that were discussed earlier.

Now, we turn our attention to simulation results that examine the sensitivity of BGP performance to policy and attack topology, when the network is subjected to peering session attacks. Fig. 22 shows the 3-D plots of AS path-lengths for the steady-state (stable) case with Policy 2 (P2) and the peak path-length for each AS-prefix pair for experiments E3P0, E3P1, and E3P2. The peak path-length here is the longest path-length recorded per AS-prefix pair during the experiment observation interval. In this set of experiments, T2-T3 BGP peering sessions are subjected to attacks with 25% probability (see experiment details stated earlier). Although not shown in the figure, the stable path-length plots for Policy 1 (P1) and no policy (P0) cases are very closely similar to that for P2 that is shown in Fig. 22. These stable path-lengths are small and four-hops long at most. However, under attack conditions, some shortest paths are suppressed due to RFD and longer paths are used instead. The corresponding path lengths (peak value under transient conditions due to attacks) are typically about 2-, 6-, and 3-times longer for the cases of experiments E3P0, E3P1, and E3P2, respectively. These increased path-lengths are sensitive to policy and topology. Due to the impact of attacks, Policy 1 makes use of fairly long alternate routes through multiple Tier 2 nodes. However, many of the same routes are not usable in the case of Policy 2 because it is much more restrictive than Policy 1. When more alternate paths are available and allowed

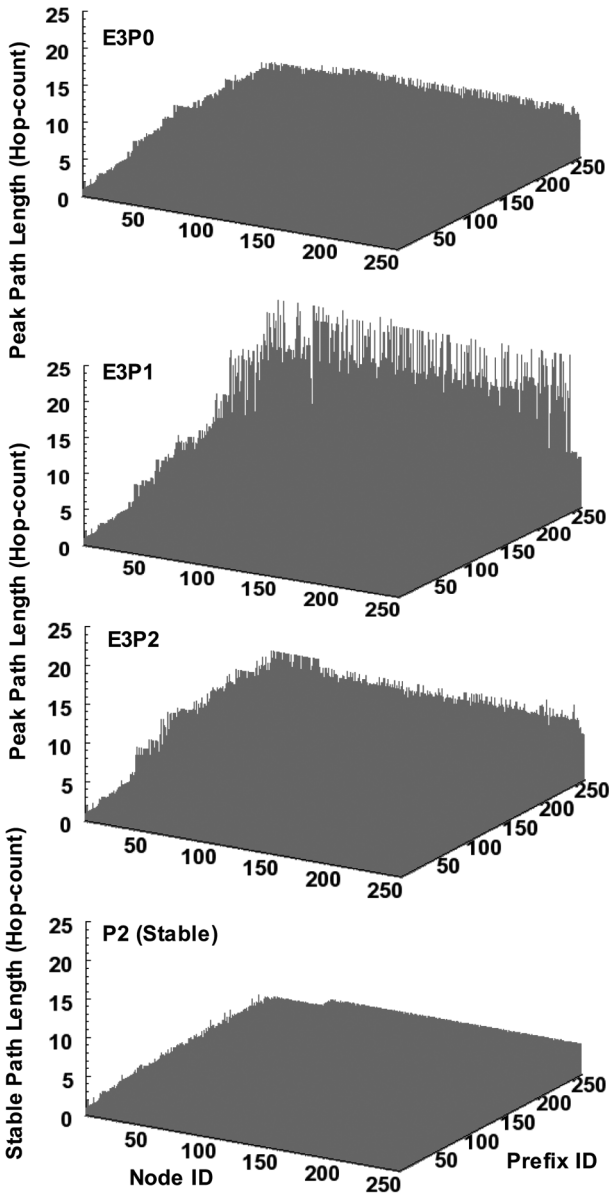


Fig. 22. AS path-length comparison for various policies (experiment set E3Px).

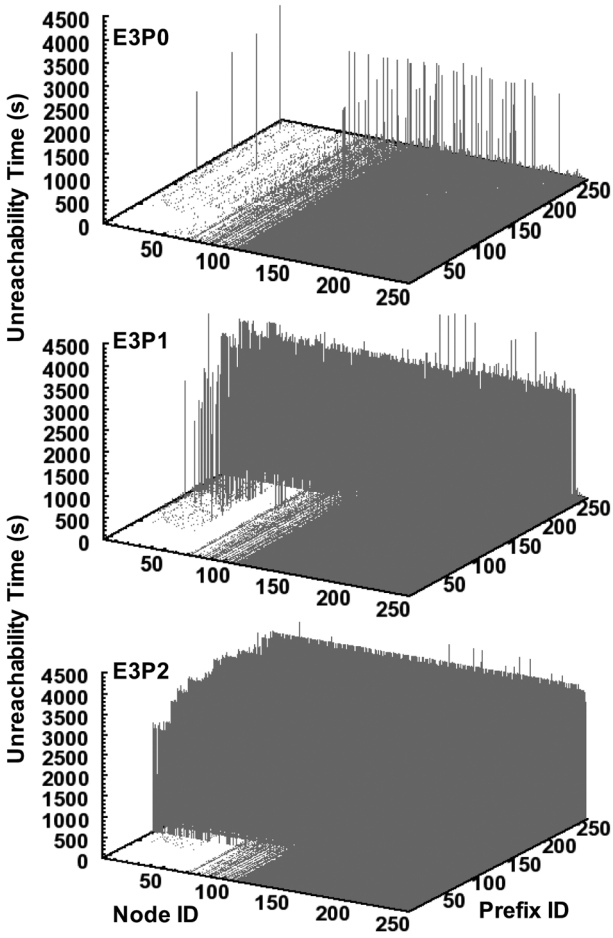


Fig. 23. Unreachability time comparison for various policies (experiment set E3Px).

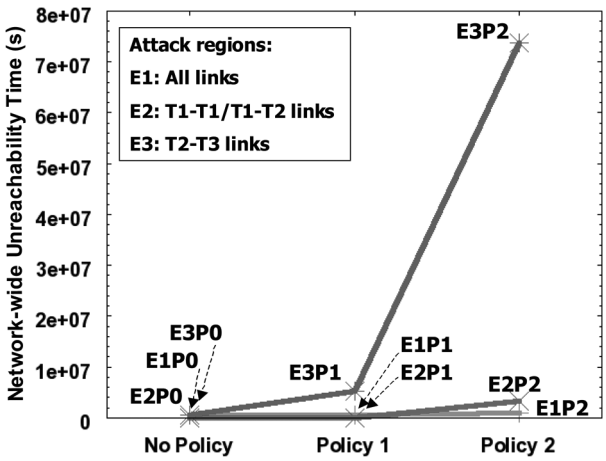


Fig. 24. Network-wide unreachability time plotted for all combinations of policy and attack topology.

for use by a policy in effect, then the AS-prefix unreachability times would be correspondingly lower. This is depicted in Fig. 23. The unreachability time increases from the least for the no Policy case to the worst for Policy 2. In other words, the performance impact is increasingly worse as we go from the least restrictive to the most restrictive policy. Fig. 24 compares the network-wide total AS-prefix unreachability time (over the experiment observation interval) for all nine experiments (see Table IV for the list of experiments). The unreachability time clearly becomes worse with use of more restrictive policies. We also observe in Fig. 24 that the same metric also dramatically varies as a function of the attack topology. We will explain this phenomenon in a moment after considering another metric of interest, namely, the number of unreachable AS-prefix pairs. This metric also behaves much the same way as the AS-prefix unreachability time in its sensitivity to policy and attack-region topology. This is evident by observing Figs. 25 and 26. Fig. 25 shows the number of unreachable AS-prefix pairs versus simulation time for all cases of experiments (E1P0 through E3P2). There is a transient period from 500 to 1000 s when attacks are in progress, and BGP sessions are reset and reestablished. Subsequently, the number of unreachable

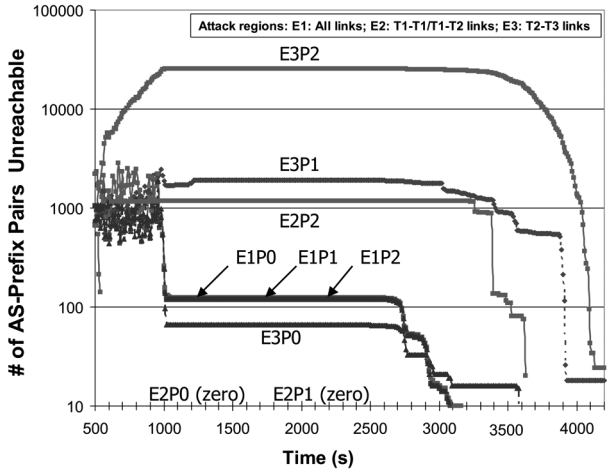


Fig. 25. Count of AS-prefix pairs unreachable as a function of time plotted for all combinations of policy and attack topology.

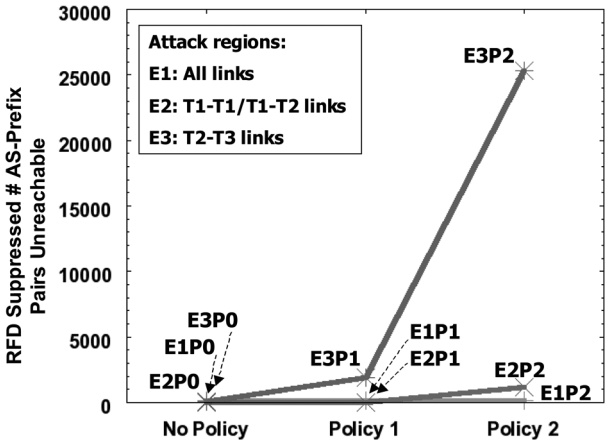


Fig. 26. Peak number of AS-prefix pairs unreachable plotted for all combinations of policy and attack topology.

AS-prefix pairs reaches a plateau when the unreachability is due to prolonged RFD suppressions. For E2P0 and E2P1, the AS-prefix unreachability counts are zero because both P0 and P1 policies allow for making ample use of alternate paths when only the richly connected portion of the network is attacked. The summary plots comparing the RFD-induced peak values of the unreachability count across all nine experiments are shown in Fig. 26.

We now provide further discussion of the results stated in the preceding paragraph (namely, those in Figs. 22–26). First, we notice that the impact of attacks on BGP performance gets worse as the route selection policy gets more restrictive. Policy 1 does not allow a path to use an intertier uplink if a downlink has already been used. Policy 2, in addition to the restriction of Policy 1, also assumes that all links in Tier 2 are private (P) links and that they can be used for transit only once in a path provided its use is immediately preceded by an uplink and immediately followed by a downlink (see Table III). With these restrictions, Policy 1 has much fewer alternate paths as compared with that for the no policy case, and Policy 2 has much fewer alternate paths than those for Policy 1. The vulnerability to RFD suppression under session attacks gets worse as the number of available alternate paths declines. It requires more effort on part of an attacker to drive an AS-prefix pair toward unreachability if the number of alternate paths for that pair is higher. As for the sensitivity to topology of the attack region, the vulnerability is lesser if the attacks are in a region where the peering connectivity is richer, and hence the available number of alternate paths is higher, which allow for better avoidance of the attacked peering links in route selection. This is the reason why the BGP routing performance is less degraded when the attack-region is T1-T1/T1-T2 (core) links as compared with the same when the attack-region is T2-T3 (edge) links. Many of the Tier 3 ASs are stub nodes that have a single-link (and single point of failure) connectivity to the rest of the network. Hence, attacking at the edges of the network (i.e., T2-T3 links) is more harmful (or more productive from an attacker's point of view) as compared with attacking the core (i.e., T1-T1/T1-T2) links (see Figs. 24–26). It is important to keep in mind here that the attacks are conducted on individual peering links and not on a BGP router as a whole.

A result that is somewhat more dramatic is seen while comparing the case of all-links attack region (E1Px set of experiments) versus the partial-network attack region consisting of T2–T3 links (E3Px set of experiments). The interesting question here is why are unreachability metrics (see Figs. 24–26) much higher for the latter case as compared with that for the former case? Note especially that in both cases each peering link is attacked with equal energy, i.e., once every 10 s with a probability of success of 25% for 50 such successive intervals. An answer for the above question emerges when we realize that RFD penalties for the prefixes reached through a peer are reinitialized when the peering session with that peer breaks and restarts. When only the T2–T3 sessions are under repeated attacks (E3Px set of experiments), then the RFD penalties build up and quickly exceed the suppression threshold at Tier 1 nodes for Tier 3 destinations via Tier 2 peers. However, if T1-T2 peering links may also come under attack (as in the case of E1Px set of experiments), then the RFD penalties at Tier 1 nodes for Tier 3 destinations will be frequently reset to zero and that helps reduce the unreachability (or outage). The attackers can exploit RFD as we have seen throughout of the results discussed in this paper, but they could also overdo it as illustrated by the comparison between E3Px versus E1Px experiments. From Figs. 24 and 26, we also observe that even attacking the more richly connected part of the network (T1-T1, T1-T2 links) causes somewhat more unreachability than attacking the entire network (e.g., compare E2P2 versus E1P2).

Based on the above results and discussion, we note that if malicious attacks are already targeted at the more vulnerable edge (T2-T3) peering links, then it is beneficial from the attackers' point of view to abstain from attacking peering links that are one or more hops away towards the core. From the network operators' perspective, it seems necessary that peering session protection mechanisms must be provided not only in the core ASs, but they must also extend to the edge ASs (i.e., customers' ASs).

## VIII. Conclusion

We have shown that routing attacks can be tuned to take advantage of the BGP protocol behavior, and thus significantly

amplify the adverse impact of those attacks on the routing infrastructure. RFD was designed to alleviate BGP processing overload and route flapping under non-malicious scenarios. However, under certain BGP peering session attack scenarios, the RFD facilitates further vulnerability in BGP by allowing severe amplification of unreachability, as well as degradation of route quality. We have presented a detailed analytical study of the impact of BGP peering session attacks that exploit the RFD. Our results have revealed that it is possible for the attackers to achieve a high probability of AS-AS and AS-prefix isolation by attacks conducted at a rate roughly equal to once per MRAI, even with a low success rate per BGP session attack. We have also shown that the RFD-based BGP vulnerability can be partially mitigated by using the BGP-GR mechanism.

We further studied the impact of BGP peering session attacks with RFD exploitation through detailed packet-level network simulations. The simulation results confirm that numerous prolonged isolations between ASs and prefixes are the result of these protocol aware attacks. We have reported results on both a canonical grid topology and a down-sampled realistic topology. The results show that there is significant sensitivity of the impact of attacks to the subtopology over which attacks are targeted. For instance, the attackers can selectively attack the peering sessions at the edges (e.g., Tier 3 to Tier 2 links), while abstaining from attacking the peering sessions on core links and still cause worse impacts on routing performance as compared with attacking over the entire network. Currently, the edge links tend to be less rigorously managed. From the network operators' perspective, it seems necessary that strong peering session protection mechanisms must be provided not only in the core ASs but they must also extend to the edge ASs (i.e., customers' ASs). Our simulation results also show that the use of route selection policy, while necessitated based on service provider relationships, can have aggravating impacts when the network is suffering malicious peering session attacks. In general, networks with more restrictive route selection policy tend to suffer more in terms of AS-prefix unreachability as compared with those with less restrictive policy under the same peering session attack scenarios.

The study reported here is a part of a larger ongoing effort at NIST to identify and characterize the risks associated with focused attacks of different types, including message spoofing/tampering attacks, on the BGP infrastructure. The work includes evaluation of the effectiveness and impact of various proposed mitigation techniques [16][31].

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for many helpful comments and suggestions, and Z. M. Mao, J. Rexford, and L. Trajkovic for information related to some details of RFD algorithms.

## REFERENCES

[1] O. Nordstrom and C. Dovrolis, "Beware of BGP attacks," *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 2, pp. 1–8, 2004.

[2] K. Butler, T. Farley, P. McDaniel, and J. Rexford, "A survey of BGP security," ATT Labs. Research, Florham Park, NJ, Tech. Rep. TD-5UGJ33, Feb. 2004. [Online]. Available: http://www.patrickmcdaniel.org/pubs/td-5ugj33.pdf, (Revised Apr. 2005)

[3] G. Goth, "Fixing BGP might be difficult—Or not so tough," *IEEE Internet Comput.*, vol. 07, no. 3, pp. 7–9, May/Jun. 2003.

[4] Z.M. Mao, R. Govindan, G. Varghese, and R.H. Katz, "Route flap damping exacerbates internet routing convergence," in *Proc. ACM SIGCOMM*, Pittsburg, PA, Aug. 2002, pp. 221–233.

[5] J. Kim, S.Y. Ko, D.M. Nicol, X.A. Dimitropoulos, and G.F. Riley, "A BGP attack against traffic engineering," in *Proc. Winter Simulation Conf.*, 2004, vol. 1.

[6] S.M. Bellovin and E.R Gansner, "Using link cuts to attack Internet routing," AT&T Labs. Research, Tech. Rep. [Online]. Available: http://www.research.att.com/smb/papers/reroute.pdf

[7] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Observation and analysis of BGP behavior under stress," in *Proc. 2nd ACM SIGCOMM Workshop on Internet Measurement Workshop*, Marseille, France, 2002, pp. 183–195.

[8] K. Zhang, X. Zhao, and S. F. Wu, "An analysis on selective dropping attack in BGP," in *Proc. IEEE Int. Perform. Comput. Commun. Conf.*, Apr. 2004, pp. 593–599.

[9] E.G. Coffman, Jr., Z. Ge, V. Misra, and D. Towsley, "Network resilience: Exploring cascading failures within BGP," in *Proc. 40th Annu. Allerton Conf. Commun., Compu. Control*, Monticello, IL, Oct. 2002.

[10] K. Varadhan, R. Govindan, and D. Estrin, "Persistent route oscillations in inter-domain routing," *Comput. Netw.*, vol. 32, no. 1, pp. 1–16, Jan. 2000.

[11] F. Gont, "ICMP attacks against TCP," IETF, Internet Draft, Dec. 2004, draft-gont-tcpm-icmp-attacks-03.txt.

[12] "Flaw Could Cripple Entire Net," Associated Press, Apr. 20, 2004. [Online]. Available: http://wired-vig.wired.com/news/technology/0,1282,63143,00.html

[13] "NISCC Vulnerability Advisory 236929: Vulnerability Issues in TCP." Apr. 20, 2004.

[14] "CERT advisory CA-2001-09: Statistical weaknesses in TCP/IP initial sequence numbers." [Online]. Available: http://www.cert.org/advisories/CA-2001-09.html. Original date May 2001, last revised Feb. 2005

[15] "Best practices guidelines: border gateway protocol," NISCC (U.K. Govt.), Apr. 2004, (see note 5 on pg. 8).

[16] D. Montgomery, K. Sriram, O. Borchert, O. Kim, and D. R. Kuhn, "Characterizing the risks and costs of BGP insecurity/security," presented at the 1st DHS Workshop on Secure Protocols for the Routing Infrastructure (DHS-SPRI), Washington, D.C., Mar. 15–16, 2005, (presentation slides available from authors upon request).

[17] S. Convery, D. Cook, and M. Franz, "An attack tree for the border gateway protocol," IETF ID, Tech. Rep., Feb. 2004. [Online]. Available: http://ietfreport.isoc.org/ids/draft-ietf-rpsec-bgpattack-00.txt

[18] Y. Rekhter and T. Li, "A border gateway protocol 4 (BGP-4)," IETF, RFC 1771, Mar. 1995.

[19] I. van Beijnum, *BGP: Building Reliable Networks with the Border Gateway Protocol*. : O'Reilly, 2002.

[20] S. Kent, C. Lynn, and K. Seo, "Secure border gateway protocol (S-BGP)," *IEEE J. Sel. Areas Commun. (Special Issue on Network Security)*, Apr. 2000.

[21] J. Ng, "Extensions to BGP to support secure origin BGP (soBGP)," IETF ID, draft-ng-sobgp-bgp-extensions-02.txt.

[22] L. Subramanian, "Listen and whisper: Security mechanisms for BGP," in *PROC First Symp. Netw. Syst. Design and Implementation*, 2004.

[23] A. Heffernan, "Protection of BGP sessions via the TCP MD5 signature option," IETF, RFC 2385, Aug. 1998.

[24] D. Pei, M. Azuma, D. Massey, and L. Zhang, "BGP-RCN: Improving convergence through root cause notification," *Comput. Netw.*, vol. 48, no. 1, pp. 175–194, May 2005.

[25] J. Rexford, A. Greenberg, G. Hjalmtysson, D.A. Maltz, A. Myers, G. Xie, J. Zhan, and H. Zhang, "Network-wide decision making: Toward a wafer-thin control plane," in *Proc. ACM SIGCOMM HotNets Workshop*, Nov. 2004.

[26] J. Cowie, A. Ogielski, B. Premore, and Y. Yuan, "Global routing instabilities during code red II and Nimda worm propagation." Sep. 2001. [Online]. Available: http://www.renesys.com/projects/bgp_instability

[27] T. Griffin, "BGP Impact of SQL Worm," 1-25-2003. Jan. 2003. [Online]. Available: http://www.research.att.com/griffin/bgp_monitor/sql_worm.html

[28] I. Dubrawsky, "Effects of worms on Internet routing stability," Jun. 2003. [Online]. Available: http://www.securityfocus.com/infocus/1702

[29] *Scalable Simulation Framework (SSFNet): Gallery of Baseline Networks*, [Online]. Available: http://www.ssfnet.org/Exchange/ gallery/index.html

[30] B.J. Premore, "An analysis of convergence properties of the border gateway protocol using discrete event simulation," Ph.D. dissertation, Dept. Comput. Sci., Dartmouth College, Hanover, NH, May 2003, Tech. Rep. TR2003-452.
[31] K. Sriram, D. Montgomery, O. Borchert, O. Kim, and D.R. Kuhn, "Study of BGP behavior under large scale attacks," Nat. Inst. Standards Technol. (NIST), Tech. Rep., 2006, in preparation.
[32] C. Villamizar, R. Chandra, and R. Govindan, "BGP route flap damping," IETF, RFC 2439, Nov. 1998.
[33] W. Shen and Lj. Trajkovic, "BGP route flap damping algorithms," in *Proc. SPECTS*, Philadelphia, PA, Jul. 2005, pp. 488–495.
[34] S.R. Sangli, Y. Rekhter, R. Fernando, J.G. Scudder, and E. Chen, "Graceful restart mechanism for BGP." IETF ID, Dec. 2004, draft-ietf-idr-restart-10.txt.
[35] L. Zhang, "internet topology," Internet Research Lab, (IRL), UCLA, Los Angeles, CA. [Online]. Available: http://irl.cs.ucla.edu/topology/, website for AS-level topology data.

**Kotikalapudi Sriram** (S'80–M'82–SM'97–F'00) received the B.S. and M.S. degrees from the Indian Institute of Technology, Kanpur, and the Ph.D. degree from Syracuse University, Syracuse, NY, all in electrical engineering.

He is currently a Senior Researcher in the Advanced Networking Technologies Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD. From 1983 to 2001, he held various positions at Bell Laboratories—the innovations arm of Lucent Technologies and formerly that of AT&T. His titles at Bell Laboratories included Consulting Member of Technical Staff (approximately top 1% of engineers in 2001) and Distinguished Member of Technical Staff. He is a contributing author and a coeditor of *Cable Modems: Current Technologies and Applications* (Piscataway, NJ: IEEE Press, 1999). He holds 16 patents and is a coinventor on ten other pending patents. He is a coinventor on a patent related to quality-of-service (QoS) management for VOIP that was recognized by the *MIT Technology Review Magazine* as one of five Killer Patents in 2004. He has published over 50 papers in various IEEE and other international journals and major conferences. His interests and responsibilities include performance modeling, network architecture, Internet routing protocol security, design of protocols and algorithms for multiservice broadband networks, voice-over-IP (VOIP), wireless access networks, IP/MPLS/ATM traffic controls, and hybrid fiber-coax networks.

**Doug Montgomery** (S'85–M'86) received the B.S. degree in computer science/mathematics from Towson State University, Towson, MD, and the M.S. degree in computer and information science from the University of Delaware, Newark.

He is the Manager of the Internetworking Technologies Research Group, Information Technology Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, MD. He provides technical leadership to research and standardization projects in areas that currently include: IPng, Internet infrastructure protection, service oriented architectures, Internet telephony technologies, advanced network metrology, and quantum information networks.

**Oliver Borchert** received the Dipl. Informatiker FH (computer science) from the University of Applied Science, Wiesbaden, Germany, in 2000.

Since 2001, he has been with the Advanced Network Technologies Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD. He was the Lead Engineer responsible for the design and implementation of the widely used GMPLS lightwave agile switching simulator (GLASS). Currently, he is developing and enhancing a detailed BGP vulnerability-testing simulator based on the scalable simulation framework (SSF). From 2000 to 2001, he was employed with CREON-LabControl AG, Germany, where he worked on the integration of digital signatures and electronic records based on the CFR-21 (Part 11) standard into the commercial products. From 1998 to 1999, he was with the Analytical Chemistry Division, NIST, where he worked on laboratory automation for data collection and analysis.

Mr. Borchert received the Second Prize for his poster presentation on "Dealing With Result Data Using a System Capability Dataset (SCD)" at the LabAutomation'99 Conference, San Diego, CA.

**Okhee Kim** received the M.S. degree in computer science from the New York Institute of Technology, Old Westbury.

Since 1988, she has been a Computer Scientist at the National Institute of Standards and Technology (NIST), Gaithersburg, MD. She has published several papers and developed working prototypes in the areas of network protocol performance measurements, simulation modeling, and networking tools. Her current research interests include simulation modeling and analysis, Internet security architecture, and routing infrastructure security.

**D. Richard Kuhn** (M'85–SM'99) received the M.S. degree in computer science from the University of Maryland, College Park, and the MBA degree from the College of William and Mary, Williamsburg, VA.

He is a Computer Scientist in the Computer Security Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD. His primary technical interests are in information security and software assurance. He is author or coauthor of more than 50 papers in these areas. From 1994 to 1995, he served as Program Manager for the Committee on Applications and Technology of the President's Information Infrastructure Task Force, and from 1996 to 1999 as Manager of the Software Quality Group at NIST. Prior to NIST, he worked as a Systems Analyst with NCR Corporation and the Johns Hopkins University Applied Physics Laboratory. He is a senior member of the IEEE.

Mr. Kuhn is co-developer of the role-based access control model used by industry and government, and for this work received a U.S. Department of Commerce Gold Medal for Scientific/Engineering Achievement in 2002 and Excellence in Technology Transfer Award from the Federal Laboratory Consortium in 1998.